

**ESTIMACION Y SELECCION DE MODELOS
ECONOMETRICOS DINAMICOS**

Antoni Espasa

**Banco de España. Servicio de Estudios.
Estudios Económicos, n.º 11 - 1978**

*Versión de la ponencia presentada al
seminario sobre «Modelos econométricos
para la economía española» que tuvo lu-
gar, en noviembre de 1976, en el Banco
de España.*

SUMARIO

	<u>Páginas</u>
RESUMEN	7
INTRODUCCIÓN	9
I. LA UTILIZACIÓN DE MÍNIMOS CUADRADOS ORDINARIOS (MCO) EN LA ESTIMACIÓN DE MODELOS SIMULTÁNEOS.....	11
II. VERIFICACIÓN Y SELECCIÓN DE MODELOS	15
II.1. DESCRIPCIÓN DEL PROBLEMA.....	15
II.2. EL PROBLEMA DEL AGOTAMIENTO DE LOS DATOS (DATA MINING) ...	16
II.3. VERIFICACIÓN SECUENCIAL DE HIPÓTESIS ...	18
II.4. MODELOS UNIECUACIONALES CON PROCESOS AUTORREGRESIVOS DE PRIMER ORDEN.....	19
II.5. MODELOS UNIECUACIONALES CON PROCESOS AUTORREGRESIVOS DE ORDEN r	22
II.6. ECUACIONES INDIVIDUALES ESTIMADAS POR VARIABLES INSTRUMENTALES	27
II.7. MODELOS SIMULTÁNEOS PEQUEÑOS	31
CONCLUSIÓN	37
BIBLIOGRAFÍA	39

RESUMEN

El problema con el que el investigador empírico se encuentra a la hora de estimar un modelo no es sólo de estimación, sino también de selección de una estructura específica dentro de un conjunto, generalmente amplio, de posibles estructuras.

Esta selección se debe realizar en base a estimaciones que tengan en cuenta hipótesis bastante generales en la definición de los modelos. En el caso de sistemas dinámicos el método de variables instrumentales autorregresivas es adecuado, mientras que la selección en base a estimaciones por mínimos cuadrados ordinarios puede conducirnos a elecciones erróneas.

La metodología presentada en la sección segunda de este trabajo está tomada de la investigación en curso que sobre este tema se está llevando a cabo en el departamento de Econometría de la «London School of Economics». El proceso de selección parte de una hipótesis mantenida general y está orientado para concluir con modelos que, siendo aptos para explicar la muestra, tengan el mayor número posible de grados de libertad.

La elección final será conveniente contrastarla, si la muestra disponible lo permite, mediante una estimación por métodos espectrales (véase Espasa y Sargan (1975)) y, en cualquier caso, debe ser compatible con la información *a priori* de la Teoría Económica. Por último, el problema del agotamiento de los datos puede aliviarse mediante la realización de tests de estabilidad muestral y post-muestral.

INTRODUCCION

El problema con que el economista empírico se encuentra a la hora de estimar un modelo macroeconómico consiste en que la Teoría Económica no le provee con una estructura específica cuyos parámetros deben estimarse mediante la utilización de las series temporales disponibles sobre las magnitudes económicas en cuestión, sino que el número de estructuras posibles que aquélla le señala es bastante amplio. Ante esta falta de precisión teórica, deben de ser los datos quienes elijan la especificación final dentro del conjunto de estructuras posibles.

Con ello tenemos que el investigador empírico no sólo ha de disponer de un método de estimación adecuado para el modelo que finalmente se estime, sino que ha de disponer de unas técnicas que le permitan estimar adecuadamente especificaciones bastante generales que la teoría económica no descarta. A partir de dichas estimaciones, el investigador tiene que verificar cuál es la especificación más simple que, entre las teóricamente aceptables, los datos no rechazan. Esta es a su vez la razón de la interdependencia de la Teoría Económica y la Econometría. Esta última necesita de aquélla para disponer de un modelo de partida, ya que los datos por sí solos no pueden identificar el universo que los genera. Pero dado que el modelo de partida no contiene una especificación única sino un conjunto —con frecuencia amplísimo— de especificaciones, los resultados econométricos, además de servir para autentificar o rechazar el modelo inicial, proveen a la Teoría de nuevos elementos con los que concretizar o reformular sus modelos.

En este trabajo nos concentramos en los problemas de selección y estimación de modelos en lo que a la formulación dinámica se refiere,

y damos por ciertos otros aspectos que definen a los mismos, como qué variables entran en ellos. Además sólo trataremos modelos lineales.

El plan en este trabajo es el siguiente. En la primera parte se tratan los peligros que supone el uso de los estimadores por mínimos cuadrados ordinarios (MCO) para este tipo de modelos y se señala como alternativa los estimadores por el método de variables instrumentales autorregresivas (VIA). Estos son consistentes bajo una serie de hipótesis más amplias, sobre las estructuras dinámicas de la parte sistemática y del término de error del modelo, que las requeridas para la consistencia de los MCO. Por ello es interesante verificar las hipótesis utilizadas en la estimación en contra de otras más restrictivas. A este tipo de cuestiones se dedica el resto de la comunicación, que empieza considerando modelos uniecuacionales para acabar con el estudio de modelos simultáneos.

I. LA UTILIZACION DE MINIMOS CUADRADOS ORDINARIOS (MCO) EN LA ESTIMACION DE MODELOS SIMULTANEOS

El problema con que nos enfrentamos al estimar los parámetros de un modelo simultáneo, o los parámetros correspondientes a una ecuación de dicho modelo, es que los MCO no sólo no son eficientes, sino que tampoco son consistentes. Es decir, incluso cuando el modelo reflejase exactamente el universo que intenta aproximar y dispusiésemos de un gran número de observaciones, el método MCO nos daría valores sesgados de los parámetros. La inconsistencia se puede corregir utilizando métodos de estimación simultánea de información limitada o, más generalmente, por el método de variables instrumentales, y la ineficiencia mediante la aplicación de métodos de información completa. Ahora bien, estos métodos de estimación presentan más problemas en su computación que los MCO, por lo que su utilización ha sido bastante restringida.

En la práctica al estimar modelos macroeconómicos bajo el supuesto de ausencia de correlación serial en las perturbaciones, los resultados obtenidos por MCO no difieren mucho, en bastantes casos, de los obtenidos con otros estimadores, lo que se ha utilizado como gran argumento para defender el uso de aquéllos. Sin embargo, como ha señalado claramente el profesor L. R. Klein, pequeñas diferencias en los coeficientes estimados pueden tener efectos serios al considerar el sistema en conjunto. Así, existe cierta evidencia (ver Klein (1969) y Seaks (1974)) de que la simulación de un modelo en base a los resultados de la estimación por MCO es peor que la obtenida en base a estimadores consistentes.

Además, si en el modelo incorporamos una hipótesis más realista que considere que el ruido u_t , que afecta al sistema en el momento t , está correlacionado con los ruidos o distorsiones que se sufrieron previamente, los métodos de estimación consistente son todavía más complicados.

Como se ha señalado en la Introducción, dada la escasa información que el econométra obtiene de la Teoría Económica sobre las propiedades de las perturbaciones aleatorias, especificación dinámica del modelo, etc., el problema con que aquél se enfrenta no es sólo de estimación, sino también de especificación. Sobre la metodología a seguir para ello el artículo de Hendry (1975a) es particularmente interesante y pasamos a considerarlo en sus rasgos generales.

El modelo de Hendry (1975a) es excesivamente sencillo y posiblemente en un modelo más complejo los sesgos serían menores. Por otra parte, como señala Sargan (ver Renton (1975), página 321), el modelo podría especificarse, alternativamente, de forma que los MCO fueran consistentes. No obstante las hipótesis necesarias para ello son menos realistas que las que supone Hendry (1975a), por lo que las conclusiones —no siempre nuevas— que éste obtiene pueden tomarse como indicativas de las que obtendríamos en modelos más complejos. El objetivo del estudio en cuestión, es investigar los efectos que ciertos errores de especificación de un modelo simultáneo y dinámico, tienen sobre la consistencia de los parámetros estimados. La estructura del modelo es tal que, dando el valor cero a ciertos coeficientes el modelo deja de ser simultáneo, dinámico o pierde la correlación serial en los errores. Ello permite estudiar los efectos de especificación incorrecta por uno o varios motivos y, aunque esto se hace de forma asintótica, es sorprendente la confirmación empírica que se obtiene de las conclusiones teóricas, con una muestra de 47 observaciones.

Sin entrar en el detalle del artículo señalemos algunos puntos de interés que en él se sugieren.

En el modelo de regresión múltiple con todas las variables explicativas exógenas, es un resultado bien conocido que si omitimos un regresor, los estimadores obtenidos para el resto de los coeficientes están, en general, sesgados (ver Theil (1971), capítulo 11). Por el contrario, si incluimos indebidamente una variable adicional en la regresión, esto no afecta a la consistencia (aunque sí a la eficiencia) del resto de los parámetros. No obstante, este resultado no es cierto si la variable que erróneamente se incluye es endógena. En tal caso los parámetros estarán, en general, sesgados incluso para grandes muestras. La consecuen-

cia es que a la vista de los resultados de esta última regresión podemos excluir, equivocadamente, en base a los estadísticos t de Student, variables que debieran incluirse e incluir indebidamente el regresor endógeno.

En la estimación MCO de un modelo de regresión con retardos endógenos y errores autorregresivos, sabemos que los estimadores que obtenemos son inconsistentes. Hendry, en base a su modelo, pone de manifiesto que la inclusión errónea de una variable exógena irrelevante no tiene un efecto adicional en la inconsistencia mencionada. Sin embargo, la adecuada introducción de una variable exógena reduce el sesgo. En ese modelo conviene observar que si la autocorrelación en los errores es negativa las inconsistencias, en valor absoluto, son mayores que cuando la autocorrelación es positiva, y el efecto dinámico en la ecuación se infraestima. Por ello y dado que la estimación de ρ , en un proceso autorregresivo de primer orden para los errores, a partir de los residuos de MCO, está sesgada hacia cero, es fácil concluir, en este caso y en base a los resultados MCO, que el retardo endógeno no es significativo y que no hay correlación serial en los errores.

Como conclusión general de los aspectos teóricos y evidencia empírica presentada en Hendry (1975a), tenemos que para ecuaciones pertenecientes a un modelo simultáneo *es peligroso seleccionar una especificación a partir de los resultados por MCO y que por el contrario es muy importante realizar la estimación en base a métodos que tengan en cuenta la simultaneidad y la correlación serial, y a partir de dichos resultados verificar la estructura propuesta por todos los modos posibles y seleccionar únicamente aquellas especificaciones que no sean rechazadas en dichos tests y correspondan a un modelo teórico aceptable.* Un método de estimación con las características mencionadas, es el propuesto por Sargan (1964) (de variables instrumentales autorregresivas, VIA) y que se puede aplicar mediante los programas RALS y GIVE (Hendry (1972 y 1973)).

II. VERIFICACION Y SELECCION DE MODELOS

II.1. DESCRIPCIÓN DEL PROBLEMA

En la sección anterior hemos apuntado que al estimar una ecuación perteneciente a un modelo simultáneo utilizando simplemente información limitada sobre el sistema, el método VIA de Sargan es adecuado, ya que corrige el sesgo debido a la determinación simultánea de la variable dependiente con algunas variables explicativas y el sesgo debido a la correlación serial de tipo autorregresivo en las perturbaciones aleatorias que afectan al sistema. Ahora bien, en la consideración de una ecuación del tipo mencionado no tenemos, en general, información exacta sobre la estructura dinámica de la parte sistemática ni sobre la estructura autorregresiva¹, AR, de los errores, por lo que dichos aspectos se han de determinar empíricamente y para ello podemos utilizar una serie de tests. Gran parte de estos tests se describen en Espasa (1973), pero allí no se señala el orden a seguir en su realización. En Anderson (1971) se establece el principio de verificar las hipótesis partiendo de la menos restrictiva a la más restrictiva, con lo que cada hipótesis se verifica sin depender de otras hipótesis que se han de verificar posteriormente. Con ello tenemos que es más probable acabar con un modelo más general que el que la realidad económica exige, que acabar con un modelo excesivamente restrictivo. La gran ventaja de

¹ La estructura de los errores puede ser también de medias móviles (MA) o del tipo autorregresivo y de medias móviles (ARMA). El ceñirnos a una estructura AR en los casos en que en realidad es MA o ARMA sólo será válido en la medida que el proceso AR utilizado aproxime a la verdadera estructura MA o ARMA. Si la parte sistemática del modelo está bien especificada se puede utilizar el método de estimación espectral expuesto en Espasa y Sargan (1975) para obtener una estimación consistente de los residuos independientemente del tipo de estructura al que éstos obedezcan.

esto es que un modelo demasiado general nos planteará problemas de eficiencia, pero no de consistencia, como ocurriría con un modelo indebidamente restrictivo. El procedimiento secuencial de Anderson tiene la propiedad de ser el uniformemente más poderoso dentro de la clase de procedimientos que fijan las probabilidades de aceptar una hipótesis menos restrictiva que la verdadera.

La exposición contenida en las secciones II.4, II.5, II.6 y II.7 se basa en una serie de trabajos, la mayor parte de ellos no publicados todavía, realizados en el Departamento de Econometría de la «London School of Economics», y principalmente en Sargan (1975 y 1976), Hendry (1975b y 1975c) y Mizon (1975 y 1976). La aportación del autor en estas secciones se reduce a la presentación sistemática de parte del material contenido en dichos trabajos ampliando aquellos puntos que se han considerado más oscuros, con material desarrollado en artículos anteriores o con comentarios personales.

Previamente dedicaremos la sección II.2 al problema del agotamiento de los datos.

II.2. EL PROBLEMA DEL AGOTAMIENTO DE LOS DATOS (DATA MINING)

Los test para la verificación de hipótesis de que hablábamos en el apartado anterior llevan consigo el problema del agotamiento de los datos, consistente en que, al someter a diversos tests el mismo conjunto de datos, los valores probabilísticos asignados a la especificación finalmente escogida pueden estar sobrevalorados. El ejemplo de Jorgenson et al. (1970), páginas 214-215, es bastante ilustrativo. En efecto, si el nivel de significación del error de tipo I (probabilidad de rechazar equivocadamente la hipótesis nula cuando es cierta) en la verificación de hipótesis es α , y se verifican p hipótesis por medio de estadísticos independientemente distribuidos, la probabilidad de rechazar la hipótesis nula al menos una vez en las p alternativas es $(1 - [1 - \alpha]^p)$, que aumentando p se puede hacer tan grande como se quiera. Este ejemplo es exagerado pues, en general, los estadísticos empleados no se distribuyen independientemente unos de otros. No obstante, el problema de agotamiento de datos es bastante grave en el campo de la economía aplicada.

Utilicemos el modelo de regresión múltiple con regresores fijos para ilustrar más claramente el problema. Supongamos que la hipótesis:

$$H_0: y = X\beta + \mu \quad [4]$$

es cierta. Supongamos también que en base a una muestra de T observaciones queremos verificar H_0 en contra de hipótesis alternativas del tipo:

$$H_i: y = X\beta + \gamma_i Z_i + v \quad [5]$$

Es decir, queremos investigar si las variables Z_1, Z_2, \dots, Z_p entran también en la explicación de y . Para ello hacemos las regresiones (con T observaciones) de y sobre X y Z_i dando a i cada vez un valor desde 1 a p , y se toma el cociente t , correspondiente al parámetro de la variable Z_i , como criterio para verificar la hipótesis nula. Supongamos que el cociente t muestral sigue una distribución t de Student y que los grados de libertad en la regresión son 20. La probabilidad de rechazar la hipótesis nula en la primera regresión (es decir, obtener un cociente t para Z_1 mayor que 2.086) es del 5 por 100, pero la probabilidad de rechazar H_0 al menos una vez en las p regresiones está comprendida entre el 0,05 y el $(1 - 0,95^p)$ por uno. Es decir, aumentando el número de variables Z_i a considerar, el límite superior que tiene la probabilidad de rechazo de la hipótesis nula, cuando es cierta, se va haciendo cada vez más próximo a cero. Dicho de otro modo, la probabilidad de que, para al menos un Z_i , el valor de t sea superior a 2.086 cuando el valor teórico del correspondiente γ_i es cero, será en general superior al 5 por 100.

Desgraciadamente dada la escasa información teórica de que disponemos para formular los modelos econométricos, y dada la escasa información estadística disponible, en la que las series temporales son cortas, y por tanto, no permiten su descomposición en p subseries (todas ellas pertenecientes al mismo modelo teórico), de forma que cada hipótesis se pudiese verificar con datos igualmente válidos pero distintos, el investigador empírico no tiene forma de evitar el problema del agotamiento de los datos.

Conviene señalar, sin embargo, que como intento de solución al problema de este apartado, se está desarrollando una literatura econométrica sobre «estimadores por contrastaciones preliminares» (preliminary test estimators), pero este tipo de estimadores sólo se han obtenido para modelos mucho más sencillos que los que nos preocupan en este documento.

De todas formas la situación no es tan pesimista, pues existen modos de confirmar la autenticidad de la bondad de las estimaciones. Esto se puede lograr principalmente verificando la estabilidad del modelo dentro de la muestra y verificando también la capacidad del modelo en predecir.

Dados los peligros, que hemos visto, que el agotamiento de datos suponen *es muy deseable que allí donde se expongan los resultados de una estimación se añadan siempre los resultados sobre los test de estabilidad en la estimación y en la predicción.*

II.3. VERIFICACIÓN SECUENCIAL DE HIPÓTESIS

Volvamos al problema de la determinación empírica de la especificación dinámica de los modelos econométricos. Señalemos que lo que se propone en las secciones siguientes es un procedimiento *ad hoc* para suplir la falta de una verdadera teoría dinámica del comportamiento económico, que nos diera información *a priori* sobre las distribuciones de retardos que aparecen en los trabajos empíricos. El desarrollo de dicha teoría (véase por ejemplo las indicaciones contenidas en Nerlove (1972)), y la posterior especificación de los modelos econométricos incorporando la información contenida en los resultados de aquélla, debería ser el camino a seguir. No obstante, los hallazgos en este sentido son escasos y la formulación dinámica de los modelos econométricos es, todavía, principalmente una cuestión empírica. Como veremos más adelante, la metodología propuesta para hallar una formulación dinámica adecuada no cubre, plenamente, todos los problemas que la falta de información teórica y las muestras disponibles (pequeñas) plantean.

La especificación dinámica de un modelo puede deberse a una estructura de retardos en la parte sistemática del mismo (estructura dinámica sistemática, EDS) y/o a una estructura dinámica de los errores (EDE). La verificación secuencial de hipótesis que se presenta en los apartados siguientes tiene como objeto proponer formulaciones de los modelos que supongan una explicación aceptable de los datos, pero con el mayor número posible de grados de libertad. El procedimiento es en cierto modo arbitrario, pues, ya hemos indicado, es simplemente una forma de contrastación estadística que viene a sustituir la falta de información teórica. Por todo ello el valor de los resultados finales ha de juzgarse en gran parte en función de opiniones *a priori* sobre la aceptabilidad económica de los resultados. No obstante, una confirmación de éstos —los cuales se obtienen a través de formulaciones de series temporales en el dominio del tiempo— se puede obtener mediante una reestimación del modelo final utilizando el análisis espectral. En efecto, en el proceso de determinación de los factores dinámicos podemos llegar a especificaciones «válidas» para el período muestral en su conjunto, pero que sean inestables dentro de él. Por ejemplo, la EDS en situaciones en las que la evolución económica es lenta puede diferir de la

correspondiente a situaciones de cambio más rápido. Este tipo de hipótesis tienen una verificación inmediata mediante el análisis espectral, estimando el modelo para bandas de frecuencias bajas (evolución lenta) y para bandas de frecuencias altas (evolución rápida) y ver si las estructuras son significativamente distintas o no.

II.4. MODELOS UNIECUACIONALES CON PROCESOS AUTORREGRESIVOS DE PRIMER ORDEN

Consideremos el modelo

$$X \underline{\gamma} = \underline{y} - Z \underline{\beta} - \alpha \underline{y}_1 = \underline{u} \quad [6]$$

en el que \underline{y} es el vector de observaciones de la variable dependiente, cuyo parámetro se ha normalizado al valor unidad, Z es la matriz de observaciones de k variables explicativas, \underline{u} es un vector de perturbaciones no observables que afectan al sistema y $\underline{\beta}$ y α son los parámetros estructurales a estimar. Un vector o una matriz con un subíndice i se refiere al vector o matriz original retardado i periodos.

Supongamos que las perturbaciones aleatorias se generan mediante un proceso autorregresivo (AR) de primer orden.

$$\underline{u} = \rho \underline{u}_1 + \underline{\varepsilon} \quad , \quad \underline{\varepsilon} = [\varepsilon_t] \quad , \quad \underline{\varepsilon} \sim N(0, I \sigma^2)^2 \quad [7]$$

El modelo definido por [6] y [7] se puede transformar como sigue

$$X \underline{\gamma} - \rho X_1 \underline{\gamma} = \underline{y} - Z \underline{\beta} - \alpha \underline{y}_1 - \rho \underline{y}_1 + \rho Z_1 \underline{\beta} + \rho \alpha \underline{y}_2 = \underline{\varepsilon} \quad [8]$$

con lo que \underline{y} depende ahora de una perturbación aleatoria del tipo «ruido blanco». A [8] le denominamos la forma transformada restringida (ya que los parámetros de [8] están sometidos a ciertas restricciones no lineales), FTR. En contraposición a [8] podemos considerar una forma transformada no restringida (FTNR)

$$X \underline{\theta} - X^0 \underline{\theta}_1 = \underline{y} - Z \underline{\delta}_1 - \delta_2 \underline{y}_1 + Z^0 \underline{\delta}_3 + \delta_4 \underline{y}_2 = \underline{v} \quad [9]$$

donde X^0 es el conjunto de variables no redundantes (es decir, variables que no son combinaciones lineales de las variables contenidas en X ; por ejemplo, el término constante, las «dummies» estacionales, etc., son variables redundantes).

Los valores de $\underline{\gamma}$ y ρ que maximizan la función de verosimilitud correspondiente a [8] cuando las observaciones iniciales son conocidas, son

² A las perturbaciones del tipo $\underline{\varepsilon}$ las denominaremos «ruido blanco».

los valores que minimizan la suma de los cuadrados de los componentes de $\underline{\varepsilon}$, S_{ε} .

Es decir, si

$$\begin{aligned}\underline{\varepsilon} &= \underline{y} - Z \underline{\beta} - (\alpha + \rho) \underline{y}_1 + \rho Z_1 \underline{\beta} + \rho \alpha \underline{y}_2 = \\ &= \underline{y} + Z \underline{\varphi}_1 + \varphi_2 \underline{y}_1 + Z_1 \underline{\varphi}_3 + \varphi_4 \underline{y}_2, \\ S_{\varepsilon} &= \underline{\varepsilon}' \underline{\varepsilon}.\end{aligned}\quad [10]$$

La minimización de esta suma de cuadrados se puede realizar bajo diferentes hipótesis que, de la menos a la más restrictiva, son:

I) Minimizar sin tener en cuenta las restricciones no lineales a que están sometidos los parámetros de [8], o lo que es equivalente minimizar la suma de los cuadrados de los residuos de [9], S_v , es decir, aplicar MCO a [9].

Si llamamos

$$\underline{\varphi}' = (\varphi'_1 \varphi'_2 \varphi'_3 \varphi'_4)$$

la hipótesis utilizada aquí es

$$\begin{bmatrix} 1 \\ \underline{\varphi} \end{bmatrix} = \begin{bmatrix} & I \\ \mu_1 & \text{---} & \\ & \text{---} & \mu_{k+1} \end{bmatrix} \underline{\gamma} \quad [11]$$

(obviamente los μ correspondientes a las variables redundantes son cero).

II) Minimizar teniendo en cuenta las restricciones paramétricas de [8], es decir, aplicando mínimos cuadrados autorregresivos (MCA). La correspondiente suma de cuadrados es S_{ε} . La hipótesis utilizada es

$$\begin{bmatrix} 1 \\ \underline{\varphi} \end{bmatrix} = \begin{bmatrix} I \\ -\rho I \end{bmatrix} \underline{\gamma} \quad [12]$$

III) Minimizar bajo la restricción de que ρ es cero, es decir, aplicando MCO a [6]. La suma de cuadrados es S_u y la hipótesis utilizada es

$$\begin{bmatrix} 1 \\ \underline{\varphi} \end{bmatrix} = \begin{bmatrix} I \\ 0 \end{bmatrix} \underline{\gamma} \quad [13]$$

Tenemos, pues, que

$$S_u \geq S_e \geq S_v \quad [14]$$

es decir, cuanto menos restrictiva es la hipótesis utilizada en la minimización, menor será la suma de cuadrados.

En esta situación hay dos tipos de hipótesis a testar:

$$H_1: \quad \mu_i = -\rho \quad i = 1, \dots, k+1$$

es decir, que la restricción autorregresiva es válida o sea que el modelo viene definido por [8] y no por [9]. El estadístico empleado puede ser

$$\lambda_1 = T \log (S_e / S_v),$$

que asintóticamente tiene una distribución $\chi^2_{(n^*)}$, donde $n^* = k - n$, donde n es el número de variables redundantes. Si λ_1 es mayor que el valor tabulado al nivel de significación deseado, rechazamos H_1 , es decir, rechazamos [8] en favor de [9]. Tenemos, pues, que no hay tal autocorrelación en las perturbaciones, sino más bien que la especificación sistemática en [6] no era correcta y necesita reformularse de acuerdo con [9]. En esta reformulación es conveniente tener en cuenta los estadísticos t y F sobre la significación individual y conjunta de los parámetros estimados en [9] para decidir qué retardos conviene incluir. Con la nueva reformulación se empezará de nuevo el proceso de verificación.

Si λ_1 es menor que el valor tabulado, [8] no se rechaza. Es decir, la reducción en la suma de cuadrados de [9] respecto a [8] no es estadísticamente significativa, en cuyo caso cogemos el modelo más simple, es decir [8]. Solamente en este caso, en que no se rechaza [8], tiene sentido pasar a verificar la siguiente hipótesis de la significación de la autocorrelación; es decir,

$$H_2: \quad \rho = 0$$

y esto se puede realizar mediante el estadístico

$$\lambda_2 = T \log (S_u / S_e)$$

que asintóticamente tiene una distribución $\chi^2_{(1)}$. Este estadístico es asintóticamente equivalente al estadístico t para el parámetro ρ en la estimación de [8]. Si λ_2 es mayor que el valor tabulado rechazamos H_2 en favor de la autocorrelación; si es menor tenemos que la reducción en

la suma de cuadrados por tener en cuenta un proceso autorregresivo no es significativamente distinta de la obtenida suponiendo errores del tipo ruido blanco y en consecuencia adoptamos esta última hipótesis por ser más sencilla.

II.5. MODELOS UNIECUACIONALES CON PROCESOS AUTORREGRESIVOS DE ORDEN r

El análisis señalado en el apartado anterior se generaliza para el caso de perturbaciones autorregresivas de orden r :

$$\underline{u} = \rho_1 \underline{u}_1 + \dots + \rho_r \underline{u}_r + \underline{\varepsilon} \quad [15]$$

con lo que si el modelo es

$$X \underline{\gamma} = \underline{u} \quad [16]$$

la FTR es

$$X \underline{\gamma} - \rho_1 X_1 \underline{\gamma} - \dots - \rho_r X_r \underline{\gamma} = \underline{\varepsilon} \quad [17]$$

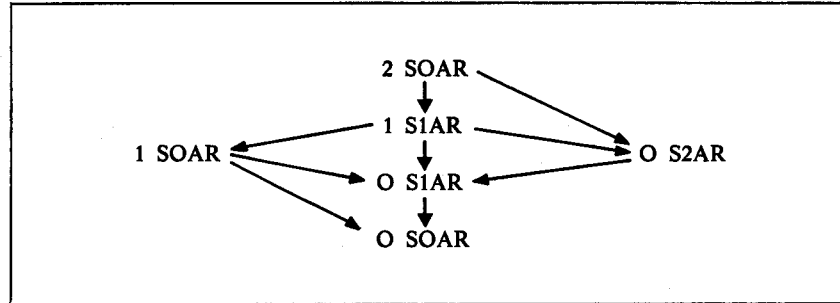
y la FTNR es

$$X \underline{\theta} - X_1^0 \underline{\theta}_1 - \dots - X_r^0 \underline{\theta}_r = \underline{v} \quad [18]$$

Si denominamos r SOAR a la hipótesis recogida en [18] de una especificación dinámica de orden r en la parte sistemática y de orden cero en las perturbaciones, y si denominamos OSr AR a la hipótesis definida en [15] y [16], es decir, de orden cero³ en la parte sistemática y de orden r en las perturbaciones, al igual que antes podemos contrastar OSr AR contra r SOAR comparando las sumas de cuadrados en ambas hipótesis mediante un estadístico λ_1 . El problema radica en que desde la hipótesis r SOAR hasta la hipótesis $OSOAR$ existen varias posibles hipótesis intermedias y todas ellas no están anidadas como ocurría antes en [14]. El número de posibilidades para $r = 2$ se puede ver en el gráfico 1, tomado de Mizon (1975), en donde las flechas van de una hipótesis menos a otra más restringida. En el gráfico se aprecia claramente que las hipótesis $OS2AR$ y $1SOAR$ no están anidadas. Para el caso general, r , tenemos que mientras la suma de los cuadrados de los

³ La terminología de orden r u orden cero en la parte sistemática se refiere al número de retardos de X , en conjunto, que aparecen en el modelo, pero en la composición de X pueden ya existir variables retardadas si existe información *a priori* para incluirlas.

GRÁFICO 1



errores correspondientes a la hipótesis r SOAR —llamémosle $S(r$ SOAR)— es menor que la correspondiente suma de cuadrados de $OSrAR$ — $S(OSrAR)$ —, es decir,

$$S(rSOAR) \leq S(OSrAR) \quad [19]$$

y a su vez

$$S(OSrAR) \leq S(OS[r-j]AR) \quad [20]$$

no se puede establecer prelación alguna entre $S([r-j]SOAR)$ y $S(OSrAR)$ para valores positivos de j y distintos de r .

Este problema se puede tratar de varias formas, una de ellas es la siguiente:

Formar con las hipótesis comprendidas entre r SOAR y $OSOAR$ todos los grupos posibles de hipótesis anidadas y aplicar la verificación secuencial en cada grupo. Con este procedimiento exhaustivo de contrastación llegaremos a diversas hipótesis posibles que se podrán verificar por medio de métodos no anidados (véase Quandt (1974)). Este procedimiento puede ser largo y su costo de computación considerable, pero además su poder estadístico puede ser pequeño. Por todo ello este procedimiento no es muy recomendable en la práctica.

Descartado este procedimiento exhaustivo las restantes alternativas se basan en limitar las hipótesis a contrastar a un conjunto anidado (o casi anidado) de ellas. El problema radica en que el test de $([r-j]SOAR)$ contra $(OSrAR)$, para $j \neq r$, no se puede encajar dentro de la secuencia que nos ha llevado a esta última. Por ello, si tenemos información suficiente para postular que la estructura sistemática del modelo se encuentra adecuadamente representada por $X\gamma$, en donde X puede contener

variables retardadas, el problema de las hipótesis no anidadas se soluciona estableciendo de antemano que la posibilidad de una estructura dinámica en la parte sistemática más allá de la contenida en X será ignorada si $(OSrAR)$ no se rechaza al ser contrastada en contra de $(rSOAR)$. Si eso es así, pasaremos a contrastar $(OSrAR)$ en contra de $(OS [r-1] AR)$. Si aquélla no se rechaza éste será el punto final. Si se rechaza en favor de $(OS [r-1] AR)$ pasaremos a contrastar ésta en contra de $(OS [r-2] AR)$ y así sucesivamente, hasta llegar a no rechazar una hipótesis (OSr^*AR) , $r \geq r^* \geq 0$. Si r^* es mayor que uno, será conveniente contrastar la hipótesis de un proceso autorregresivo de orden r^* , (OSr^*AR) , en contra de un proceso autorregresivo simple de orden r^* definido por

$$\underline{u} = \rho^* \underline{u}_r + \underline{\varepsilon} \quad [21]$$

Con datos trimestrales este test será especialmente importante para $r^* = 4$ (véase Wallis (1972)).

Así, pues, cuando la EDS no se pone en duda, la verificación secuencial es aplicable para determinar el orden del proceso autorregresivo. Sin embargo, como hemos señalado antes, la hipótesis de procesos AR o tiene por qué ser cierta, pudiendo los errores generarse por otro tipo de proceso ARMA, en cuyo caso los estimadores obtenidos para $\underline{\gamma}$ en el modelo OSr^*AR serían inconsistentes. Por ello si la EDS es correcta, es preferible estimarla bajo la hipótesis de que los errores observan una estructura estacionaria general (EEG), sin referirse a ninguna parametrización finita concreta. El método de estimación para tal caso, aplicando el análisis espectral, se describe en Espasa (1976). Con este método los errores estimados son consistentes independientemente del proceso ARMA que sigan y pueden utilizarse, aplicándoles el método Box-Jenkins (1971), para determinar qué tipo ARMA siguen, y estimar, finalmente, $\underline{\gamma}$ bajo tal hipótesis.

La verificación secuencial de los párrafos anteriores surgía al no rechazar $(OSrAR)$ en favor de $(rSOAR)$. Si por el contrario se rechaza $(OSrAR)$, tenemos que los datos no soportan la hipótesis *a priori* sobre la EDS, y el modelo debe reformularse. Al rechazar $(OSrAR)$ en favor de $(rSOAR)$, no es aconsejable, por las razones antes indicadas en II.4, contrastar esta hipótesis en contra de otras contenidas en otras sendas.

La elección de qué variables escoger entre las que forman el conjunto $\{X_1, X_2, \dots, X_r\}$ se puede realizar mediante tests F . Conviene advertir que estos tests F no están animados con el test inicial, con lo que el poder estadístico de este procedimiento se verá también afectado. Esto

parece inevitable, pero al menos el procedimiento seguido habrá servido para indicarnos que los datos no respaldaban la hipótesis inicial sobre la parte sistemática de nuestro modelo. Una vez reestructurado éste, será conveniente empezar el proceso de nuevo.

Una tercera alternativa, que en la práctica es la más recomendable y es la adoptada en Sargan (1975) y Mizon (1976), consiste en un proceso de dos etapas. Si tenemos un modelo que en su forma no restringida es

$$X \underline{\theta}_m(L) = \underline{y} \quad [22]$$

donde $\underline{\theta}_m(L)$ es un vector polinomial de orden m en el operador de retardos L , de la forma:

$$\underline{\theta}_m(L) = \underline{\theta}_0 + \underline{\theta}_1 L + \dots + \underline{\theta}_m L^m, \quad [23]$$

la primera etapa consiste en determinar m y en la segunda, condicional al valor \bar{m} determinado en la primera, se verifica si $\underline{\theta}_{\bar{m}}(L)$ se puede factorizar de la forma

$$\underline{\theta}_{\bar{m}}(L) = \rho_r(L) \gamma_{(\bar{m}-r)}(L) \quad [24]$$

donde $\gamma_{(\bar{m}-r)}(L)$ representa una EDS de orden $(\bar{m} - r)$ y $\rho_r(L)$ una EDE de orden r . La hipótesis más general es la de $r = 0$ y la más restrictiva la de $r = \bar{m}$.

Para determinar el valor de m en la primera etapa, se formula [22] con un valor $m = m^*$ y, mediante un test F , se verifica si $\underline{\theta}_{m^*} = 0$, si no se rechaza se verifica si $\underline{\theta}_{m^*-1} = 0$ y así, sucesivamente, hasta rechazar $\underline{\theta}_{\bar{m}} = 0$. El valor m^* dependerá del modelo en cuestión y de la muestra disponible, ya que los problemas de multicolinealidad pueden hacer [22] inestimable. No obstante, como señala Mizon (1976), m^* debe ser un valor que pueda recoger adecuadamente los efectos de estacionalidad y tendencia. Así, con datos trimestrales, un valor mínimo para m^* debiera ser 5.

Esta forma de determinar \bar{m} impone que las estructuras de retardos en cada uno de los componentes de X sea del mismo orden, mientras que en la práctica puede existir información *a priori* que indique que dichos órdenes son diferentes para las distintas variables de la regresión. El problema radica en que investigar los \bar{m} , correspondientes a cada variable no se puede realizar mediante una verificación secuencial que tenga las propiedades señaladas en Anderson (1971), ya que todas las hipó-

tesis posibles no están anidadas y el orden que se siga en la verificación será arbitrario. No obstante, para cada variable X_i la investigación de su correspondiente \bar{m}_i sí que constituye un tipo secuencial de los señalados en Anderson (1971). En la práctica, a pesar de la arbitrariedad mencionada, la determinación independiente de cada \bar{m}_i es recomendable.

Una vez determinados \bar{m} o los diferentes \bar{m}_i , en cuyo caso $\bar{m} = \min(\bar{m}_i)$, pasamos a verificar si el polinomio $\vartheta_{\bar{m}}(L)$ se puede descomponer según [24] dando a r los valores de 0 a \bar{m} . La hipótesis de $r = 0$ es la menos restrictiva, en la que toda la estructura dinámica se atribuye a la parte sistemática. Así, pues, contrastaremos esta hipótesis en contra de la correspondiente a $r = 1$. Si $r = 1$ no se rechaza la contrastaremos con $r = 2$ y así, sucesivamente, hasta no rechazar un valor de r en contra de otro superior en una unidad o hasta llegar a $r = \bar{m}$, que es la hipótesis más restrictiva de las definidas en [24]. En ella la estructura dinámica se inserta en las perturbaciones estocásticas del modelo. El estadístico empleado en la contrastación de hipótesis de esta segunda etapa puede ser el cociente de la máxima verosimilitud. No obstante el empleo de este estadístico puede presentar problemas cuando el proceso AR sea de orden par con todas las raíces complejas. Por ello Mizon (1976) sugiere contrastar la factorización de $\vartheta_{\bar{m}}(L)$ para valores de $r = 0, 2, 4, \dots$ y si no se rechaza un valor de $r = 2p$ pasaremos a contrastar la hipótesis correspondiente a $r = 2p$ con la correspondiente a $r = 2p + 1$.

Como estrategia para este procedimiento, Mizon (1976) sugiere elegir un m^* bastante grande (con lo que la hipótesis mantenida es bastante general) y un nivel de significación pequeño (entre el 1,0 y 0,5 por 100) en las contrastaciones próximas a la hipótesis mantenida (HM). Con ello, si la hipótesis mantenida es cierta todavía puede ser detectado por los datos. Se requerirá que el valor del estadístico empleado sea verdaderamente alto para superar un valor crítico alto correspondiente a un nivel de significación pequeño. Por otra parte si no es necesaria una hipótesis muy general la probabilidad de aceptarla (es decir, el nivel de significación del test) es pequeña.

Los tests descritos son válidos asintóticamente, es decir, para grandes muestras. Sin embargo, los vamos a aplicar a modelos estimados con pequeñas muestras. Esta discrepancia entre la teoría y la práctica se puede paliar aplicando una corrección de grados de libertad a los estadísticos sugeridos arriba. Así, si llamamos λ_R al cociente de la máxima verosimilitud entre dos hipótesis, en las que la más restrictiva contiene R restricciones sobre la más amplia, la corrección que sugiere Mizon

(1976) consiste en emplear el estadístico

$$\lambda_R^* = \frac{T^*}{T} \lambda_R$$

donde $T^* = T - k - 1 + R/2$, siendo k el número de parámetros en la hipótesis amplia, y comparar λ_R^* con el valor tabulado para χ^2 con R grados de libertad.

Otra forma de adaptar el contraste de hipótesis a pequeñas muestras es obteniendo una aproximación de la distribución de los estadísticos empleados en pequeñas muestras. Este es el procedimiento seguido en Sargan (1976).

Hemos señalado que el estadístico empleado puede ser el del cociente de la máxima verosimilitud, que implica estimar el modelo para cada hipótesis, pero que es fácil de calcular. Alternativamente se pueden obtener estadísticos basados en el principio de Wald, que sólo requieren estimar la hipótesis más general, pero que suponen una complicación considerable en su cálculo. Una programación del estadístico de Wald para este tipo de problemas se está realizando en la London School of Economics (véase Sargan (1975)).

II.6. ECUACIONES INDIVIDUALES ESTIMADAS POR VARIABLES INSTRUMENTALES

Los tests descritos en la sección anterior se generalizan para estimaciones por el método de variables instrumentales (VI) y por tanto para estimaciones por mínimos cuadrados bietápicos (MCB). Ahora, suponiendo que el número de instrumentos disponibles es superior al de regresores en la ecuación, la suma de los cuadrados en vez de ser del tipo $\underline{\gamma}' X' X \underline{\gamma}$, como ocurría en la sección anterior, serán del tipo $\underline{\gamma}' X' \bar{Z} (\bar{Z}' \bar{Z})^{-1} \bar{Z}' X \underline{\gamma} = \underline{\gamma}' R \underline{\gamma}$, donde \bar{Z} es el conjunto de instrumentos disponibles, que incluirá aquellos regresores en X que actúen de instrumentos de sí mismos. En el caso de MCB, \bar{Z} es el conjunto de variables predeterminadas que aparecen en el sistema.

Asimismo, los estadísticos empleados en los tests no provienen de cocientes de funciones de verosimilitud, sino que su justificación es la siguiente (véase Sargan (1958)). Supongamos la ecuación

$$\underline{y} - X^* \underline{\gamma}_1 = X \underline{\gamma} = \underline{u} \quad \underline{u} \sim N(O, \sigma^2 I) \quad [25]$$

en la que, además de la variable normalizada, existen otra, u otras, que se determinan en otras ecuaciones del sistema.

Formemos la variable

$$\underline{w} = \frac{\bar{Z}' u}{\sqrt{T}}$$

Si \bar{Z} constituye un conjunto adecuado de N variables instrumentales, es decir si no están correlacionadas con el término de error, de acuerdo con los resultados de Mann y Wald (1943), tenemos que

$$\underline{w} \underset{a}{\sim} N(O, M \sigma^2),$$

donde

$$M = p \lim \left(\frac{\bar{Z}' \bar{Z}}{T} \right),$$

con lo que

$$\frac{\underline{w}' M^{-1} \underline{w}}{\sigma^2} \sim \chi^2_{(N)}. \quad [26]$$

En Sargan (1958) se demuestra que

$$\begin{aligned} \lambda_o &= \frac{Q}{S_u^2} = \frac{\hat{u}' \bar{Z} (\bar{Z}' \bar{Z})^{-1} \bar{Z}' \hat{u}}{S_u^2} = \\ &= T \left(\frac{\hat{\gamma}' X' \bar{Z} (\bar{Z}' \bar{Z})^{-1} \bar{Z}' \hat{\gamma} X}{\hat{\gamma}' X' X \hat{\gamma}} \right) \underset{a}{\sim} \chi^2_{(N-k^*)} \end{aligned} \quad [27]$$

donde $S_u^2 = \hat{u}' \hat{u} / T$, $\hat{\gamma}$ se refiere a los valores estimados por V.I. y k^* es el número de parámetros a estimar en [25].

El estadístico λ_o sirve para contrastar si «la hipótesis de que entre las variables sugeridas existe una relación estructural con un término de error independiente de todos los instrumentos» (Sargan (1958), página 404), ya que, bajo dicha hipótesis, λ_o tiene una distribución asintótica $\chi^2_{(N-k^*)}$. Este test es asintóticamente equivalente al desarrollado por Hood y Koopmans (1953), págs. 178-183, para la estimación por el método de la máxima verosimilitud de información limitada (MVIL). En el caso de MCB, λ_o es un test sobre la especificación de la ecuación, y

sobre si las variables definidas como exógenas son realmente exógenas. Tenemos, pues, que en los casos en que la ecuación está sobreidentificada⁴ ($[N - k^*] > 0$), además de los tests correspondientes a los desarrollados en la sección anterior, y que comentamos más adelante, tenemos este test adicional sobre la especificación y/o sobre la sobreidentificación de la ecuación. En este test, la hipótesis alternativa no la constituye una especificación única, sino un conjunto de especificaciones que responden al tipo general de estructura establecido en la definición de [25] y a la definición de qué variables son dependientes y cuáles son predeterminadas o instrumentales (Hood y Koopmans [1953], pág. 179).

Este test es asintóticamente válido para modelos no lineales (véase Sargan (1959), pág. 29). Supongamos que en [25] el término de error no es ruido blanco, sino que sigue un proceso AR de orden r , y denominemos [25'] a la ecuación así definida. Si transformamos la matriz X de acuerdo con el filtro autorregresivo del término de error $-\rho(L)$ y llamamos X^+ a la matriz transformada (que dependerá de X y de los parámetros autorregresivos) tenemos que la forma transformada restringida de [25'] será:

$$X^+ \gamma = X \rho(L) \gamma = \varepsilon, \quad [28]$$

que no es lineal. Si llamamos

$$X^r = (X, X_1, \dots, X_r)$$

tenemos que el numerador de λ_0 para este caso será:

$$Q^r = \hat{f}(\gamma, \rho) X^{r'} \bar{Z} (\bar{Z}' \bar{Z})^{-1} \bar{Z}' X^r \hat{f}(\gamma, \rho), \quad [29]$$

donde si $\rho' = (\rho_1, \dots, \rho_r)$, $\hat{f}(\gamma, \rho)$ es función de los parámetros en γ y ρ .

El denominador será

$$S_\varepsilon^2 = \hat{f}(\hat{\gamma}, \hat{\rho}) X^{r'} X^r \hat{f}(\hat{\gamma}, \hat{\rho}) / T$$

y los grados de libertad del test serán el número de variables instrumentales menos el número de parámetros estimados, es decir, $N - (k^* + r)$.

La forma transformada no restringida se puede escribir

$$X^0 \delta = v, \quad [30]$$

⁴ Si la ecuación está justamente identificada, es decir, si el grado de sobreidentificación $(N - k^*)$ es cero, $(X' \bar{Z})^{-1}$ existe y el numerador de λ_0 es cero.

donde X^o se ha formado a partir de X' suprimiendo las variables redundantes.

El correspondiente test λ_o para el modelo [30] viene dado por

$$\frac{Q^o}{S_o^2} = \frac{\hat{\xi}' X^{o'} \bar{Z} (\bar{Z}' \bar{Z})^{-1} \bar{Z}' X^o \hat{\xi}}{(\hat{\xi}' X^{o'} X^o \hat{\xi})/T}$$

La generalización de los tests de la sección II.4 para el caso de variables instrumentales se puede ver en Hendry (1975b), y se basa en el resultado de Sargan (1958), en el que

$$\frac{(\hat{\gamma} - \gamma)' X' \bar{Z} (\bar{Z}' \bar{Z})^{-1} \bar{Z}' X (\hat{\gamma} - \gamma)}{\sigma^2} \sim_a \chi^2_{(k^*)} \quad [31]$$

y es independiente de [26].

Consideramos aquí procesos AR de primer orden, por lo que en lo sucesivo [28] y [30] se definen para $r = 1$.

Los criterios propuestos por Hendry para contrastar la validez de las restricciones autorregresivas y, en su caso, la de $\rho = 0$, se basan en lo siguiente. Utilizando [31] tenemos que bajo la hipótesis nula de que el modelo [25] es cierto

$$\lambda = \left(\frac{Q}{\sigma^2} - \frac{Q^o}{\sigma^2} \right) \sim_a \chi^2_{(k^{**})} \quad [32]$$

Conociendo σ^2 , λ se podría utilizar para contrastar si las k^{**} variables añadidas en [30] son conjuntamente significativas. En la práctica, Hendry propone sustituir σ^2 por su estimador bajo la hipótesis más amplia y corregido de grados de libertad. Si ese estimador tiene una distribución χ^2 entonces la distribución asintótica de

$$\lambda^c = \frac{Q - Q^o}{k^{**} \bar{S}_v^2}$$

donde \bar{S}_v^2 es S_v^2 corregido de grados de libertad, se puede aproximar asintóticamente por $F_{T-k^{**}-k^{**}}^{k^{**}}$ y, para valores de λ^c superiores a los tabulados para F , al nivel de significación requerido, rechazamos [25] en favor de [30].

De forma similar la verificación de las restricciones autorregresivas se puede hacer mediante el estadístico

$$\lambda_1 = \frac{Q^r - Q^o}{(k^{**}-1) \bar{S}_v^2} \quad [33]$$

comparándolo con el valor tabulado de $F_{T-k^{**}-k^{**}}^{k^{**}-1}$ y rechazando [28] en favor de [30] para valores de λ_1 superiores a los tabulados.

Si la autocorrelación serial no se rechaza podemos contrastar si $\rho = 0$, es decir, podemos pasar a verificar [25] en contra de [28]. El estadístico para ello es

$$\lambda_2 = \frac{Q - Q^r}{S_e^{2c}} \quad [34]$$

que lo compararemos con el valor tabulado para $F_{T-k^{**}-1}^1$.

Como señala Hendry (1974), pág. 564, para que en las estimaciones por VI sea posible la comparación de las ecuaciones [25], [28] y [30] se necesita que todos los regresores predeterminados que aparecen en [30] se usen como instrumentos; también en la estimación de [25] y [28], además de las variables predeterminadas de la forma reducida que se hayan seleccionado como instrumentos.

Por último señalemos que como advierte Sargan (1958), sección 11, la utilización en la estimación por VI, de estos tests asintóticos como una aproximación para el caso de pequeñas muestras, puede ser bastante pobre. En estos casos la orientación de Sargan (1976), consistente en aproximar la distribución de los estadísticos para muestras pequeñas, es bastante prometedora.

II.7. MODELOS SIMULTÁNEOS PEQUEÑOS

El procedimiento de selección señalado en las secciones anteriores se puede generalizar para tratar modelos simultáneos. En la práctica el análisis que se propone en este epígrafe sólo será aplicable a modelos muy pequeños con estructuras dinámicas cortas. Por ello sólo consideraremos procesos autorregresivos de primer orden. La exposición de este apartado está basada en Hendry (1975c).

Consideremos el modelo simultáneo lineal

$$AX' = BY' + C^* Z^{**} = BY' + CZ' + DY'_1 = U' \quad [35]$$

en el que Y y Z^* son las matrices de observaciones de las variables endógenas y predeterminadas respectivamente. Y_1 es la matriz Y retardada un período. B — $(n \times n)$ —, C — $(n \times k)$ — y D — $(n \times n)$ — son las matrices de parámetros a estimar en las que suponemos que existen suficientes restricciones cero para que el modelo esté identificado. Asimismo suponemos que B es una matriz no singular, pues de lo contrario [35] no representaría una teoría de formación conjunta de las variables dependientes. U es la matriz de perturbaciones aleatorias no observables, sobre las que si \underline{u}_t es un vector fila en U suponemos que

$$\underline{u}_t = R \underline{u}_{t-1} + \underline{\mu}_t, \quad \underline{\mu}_t \sim N(O, \Omega). \quad [36]$$

La forma reducida (FR) de [35] viene dada por

$$\text{FR1} \quad Y' = \Pi^* Z'^* + V' = \Pi_1 Z' + \Pi_2 Y'_1 + V', \quad [37]$$

donde

$$\Pi^* = -B^{-1}C^*, \quad \Pi_1 = -B^{-1}C, \quad \Pi_2 = -B^{-1}D,$$

$$V' = (\underline{v}'_t) = B^{-1}U'$$

$$\underline{v}_t = S \underline{v}_{t-1} + \underline{\eta}_t.$$

La matriz Π^* está sometida a una serie de restricciones consecuencia de las restricciones cero de B y C^* .

Al igual que en secciones anteriores podemos transformar [37] para eliminar la autocorrelación. Así tenemos:

$$\text{FR2} \quad Y' = \Pi_1 Z' + (\Pi_2 + S) Y'_1 - S \Pi_1 Z'_1 - S \Pi_2 Y'_2 + \mathfrak{J} \quad [38]$$

Por otra parte la FR sin las restricciones de sobreidentificación se puede representar de la siguiente forma:

$$\text{FR3} \quad Y' = P^* Z'^* + W' = P_1 Z' + P_2 Y'_1 + W' \quad [39]$$

donde

$$W' = S W'_1 + \Xi'.$$

Con la transformación AR de W' se obtiene:

$$\text{FR4} \quad Y' = P_1 Z' + (P_2 + S) Y'_1 - S P_1 Z'_1 - S P_2 Y'_2 + \Xi' \quad [40]$$

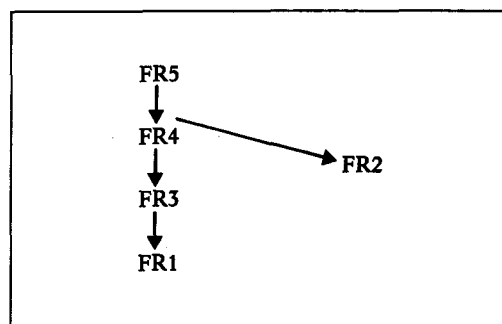
y si en [40] eliminamos las restricciones que la autocorrelación supone podemos escribir:

$$\text{FR5} \quad Y' = Q_1 Z' + Q_2 Y'_1 + Q_3 Z'_1 + Q_4 Y'_2 + E', \quad [41]$$

donde Z'_1 contiene solamente los regresores no redundantes de Z_1 .

Tenemos, pues, cinco posibles formas reducidas que están relacionadas según se muestra en el gráfico 2.

GRÁFICO 2



En estas formas reducidas se recogen tres tipos de restricciones:

- I) Las restricciones autorregresivas no lineales.
- II) Parámetros autorregresivos cero.
- III) Restricciones de sobreidentificación.

Lógicamente las restricciones de sobreidentificación deberán contrastarse una vez elegida la especificación dinámica del modelo y no antes. Es decir, antes de intentar sobreidentificar el modelo imponiendo restricciones en su estructura, es conveniente especificar las características dinámicas de ésta.

Llamemos L_x al logaritmo de la función de verosimilitud cuando el error de la FR es x . Una verificación secuencial de las formas reducidas FR1 a FR5, yendo de la menos a la más restrictiva y contrastando primero los aspectos dinámicos y en segundo lugar las condiciones de sobreidentificación, es la siguiente.

A) Verificaremos las restricciones autorregresivas. Es decir, contrastaremos

$$H_0: \quad \text{FR4}$$

en contra de

$$H_1: \text{FR5},$$

mediante el estadístico del cociente de máxima verosimilitud

$$\lambda_1 = 2 (L_E - L_{\Xi}) \underset{a}{\sim} \chi^2_{(m)}$$

donde si m° es el número de variables en Z° , $m = m^\circ \times n$. Valores de λ_1 superiores al tabulado indican que el incremento en el logaritmo de la función de verosimilitud al relajar las restricciones autorregresivas es estadísticamente significativo y en consecuencia elegimos FR5. Valores inferiores al tabulado indican que el incremento no es significativo, por lo que elegimos el modelo más restrictivo, es decir, FR4.

Si FR4 se rechaza en favor de FR5 tenemos que los datos están indicando que la estructura inicial de [35] debe modificarse incluyendo variables de las contenidas en Z°_1 e Y_2 . Al igual que en secciones anteriores los tests para decidir qué elementos de Z°_1 e Y_2 deben incorporarse no están anidados con el realizado para contrastar FR4.

B) Si en A) no rechazamos FR4 pasaremos a contrastar FR3 en contra de FR4. Es decir, pasamos a verificar si los parámetros autorregresivos son cero. El estadístico será

$$\lambda_2 = 2 (L_{\Xi} - Y_W) \underset{a}{\sim} \chi^2_{(n^2)}$$

y rechazamos FR3 para valores altos de λ_2 . De lo contrario elegimos FR3 por ser el modelo más simple.

C) Si FR3 se rechazó en B) podemos pasar a verificar las restricciones de sobreidentificación, contrastando FR4 con FR2 mediante el estadístico

$$\lambda_3 = 2 (L_{\Xi} - L_{\mathfrak{J}}) \underset{a}{\sim} \chi^2_{(N)}$$

donde N es el número de restricciones de sobreidentificación.

Si FR3 no se rechazó el test de sobreidentificación se hará contrastando FR3 y FR1 mediante

$$\lambda_4 = 2 (L_W - L_V) \underset{a}{\sim} \chi^2_{(N)}.$$

Si las condiciones de sobreidentificación se rechazan los datos indican que las restricciones cero impuestas no son compatibles con la muestra y que necesitan revisarse.

Al aplicar estos tests a modelos estimados con muestras pequeñas, necesitaremos aplicar una corrección por grados de libertad similar a la indicada en la sección II.5.

Para un programa que realice el tipo de estimaciones y cálculos que hemos discutido en esta sección, véase Hendry y Tremagne (1973).

CONCLUSION

La selección de modelos se debe realizar en base a estimaciones que tengan en cuenta hipótesis bastante generales en la definición de aquéllos. En el caso de modelos simultáneos dinámicos el método de variables instrumentales autorregresivas es adecuado, mientras que la selección en base a estimaciones por mínimos cuadrados ordinarios puede conducirnos a elecciones erróneas.

La metodología presentada en la sección segunda para realizar la selección de modelos parte de una hipótesis mantenida general y está orientada para concluir con modelos que, siendo aptos para explicar la muestra, tengan el mayor número posible de grados de libertad. La elección final será conveniente contrastarla mediante una estimación por métodos espectrales y, en cualquier caso, debe ser compatible con la información *a priori* de la Teoría Económica.

Por último dado que el problema del agotamiento de los datos puede ser grave en la selección de modelos debe considerar también la estabilidad muestral y post-muestral de éstos.

BIBLIOGRAFIA

- ANDERSON, T. V., 1971: *The Statistical Analysis of Time Series*, John Wiley.
- BOX, G. E. P., y G. M. JENKINS, 1971: *Time Series Analysis Forecasting and Control*, Holden-Day, 2.ª impresión.
- ESPASA, A., 1973: «Notas sobre la correlación serial en modelos econométricos uniecuacionales», *Anales de Economía*, enero-marzo, 95-115.
- , 1976: «The estimation of the multiple regression model with stationary errors and lagged endogenous variables». Ponencia presentada al Congreso Europeo de la Econometric Society, agosto, Helsinki.
- , y J. D. SARGAN, 1975: «The spectral estimation of simultaneous equation systems with lagged endogenous variables». Ponencia presentada al Congreso Mundial de la Econometric Society, Toronto. Una versión revisada de dicha ponencia ha sido publicada en *International Economic Review*, octubre 1977.
- GOLDFELD, S. M., y R. E. QUANDT, 1972: *Non linear Methods in Econometrics*, North-Holland.
- HENDRY, D. F., 1972: «Unser's manual for RALS, the estimation of linear equations with lagged dependent variables and autoregressive errors», *London School of Economics*. Trabajo no publicado.
- , 1973: «User's manual for GIVE general instrumental variable estimation of linear equations with lagged dependent variables and first order autoregressive errors», *London School of Economics*. Trabajo no publicado.
- , 1974: «Stochastic Specification in an aggregate demand model of the United Kingdom», *Econometrica*, vol. 42, mayo, n.º 3, 559-578.
- , 1975a: «The Consequences of Mis-specification of Dynamic Structure, Autocorrelation, and simultaneity in a Simple Model with an Application to the Demand for Imports». Capítulo 11 en Renton (1975).
- , 1975b: «On Testing Dynamic Specification», *London School of Economics*. Trabajo no publicado.
- , 1975c: «Testing Dynamic Specification in a model of Building Society Behaviour». Ponencia presentada en el «Workshop» de verano de la Econometric Society, Lovaina.

- , y A. R. TREMAYNE, 1973: «Estimating Systems of Dynamic Reduced Form Equations with Vector Autoregressive errors», *London School of Economics*. Trabajo no publicado.
- HOOD, W. C., y T. C. KOOPMANS, 1953: *Studies in Econometric Methods*, John Wiley.
- KLEIN, L. R., 1969: «Estimation of interdependent Systems in Macro-econometrics», *Econometrica*, vol. XXXVII, 171-192.
- JORGENSEN, D. W.; J. HUNTER, y M. I. NADIRI, 1970: «The predictive performance of econometric models of quarterly investment behaviour», *Econometrica*, vol. 38, n.º 2, 213-224.
- MANN, H. B., y A. WALD, 1943: «On the Statistical Treatment of linear stochastic difference equations», *Econometrica*, julio.
- MIZON, G., 1975: «Model Selection procedures in Dynamic Models». Ponencia presentada al Congreso Mundial de la Econometric Society, agosto, Toronto.
- , 1976: «Testing dynamic specification in a model of the demand for consumer durables in Canada». Ponencia presentada al Congreso Europeo de Econometric Society, agosto, Helsinki.
- NERLOVE, M., 1972: «Lags in Economic Behavior», *Econometrica*, vol. 40, n.º 2, marzo, 221-252.
- QUANDT, R. E., 1974: «A comparison of methods for testing nonnested hypotheses», *Review of Economics and Statistics*, febrero, 92-99.
- RENTON, G. A., 1975: *Modeling the Economy*, Heineman Educational Books, Londres.
- SARGAN, J. D., 1958: «The estimation of Economic relationships using instrumental variables», *Econometrica*, vol. 26, 393-415.
- , 1959: «The estimation of relationships with autocorrelated residuals by the use of instrumental variables», *Journal of the Royal Statistical Society, Serie B*, 91-105.
- , 1964: «Wages and Prices in the United Kingdom: A study in Econometric Methodology», *Econometric Analysis for National Economic Planning* (P. E. Hart, G. Millos y J. K. Whitaker), Butterworths (Colston Papers, vol. 16).
- , 1975: «A suggested technique for computing approximations to Wald Criteria with application to testing dynamic specifications», *Social Science Research Council, London School of Economics. Econometrics Programme*, «discussion paper», n.º A2, noviembre.
- , 1976: «Some approximation to the distribution of econometric criteria which are asymptotically distributed as chi square». Ponencia presentada al Congreso Europeo de la Econometric Society, agosto, Helsinki.
- SEAKS, T. G., 1974: «Simulations with Econometric Models and Alternative Methods of Estimation», *Southern Economic Journal*, vol. 41, n.º 1, 1-9.
- THEIL, H., 1971: *Principles of Econometrics*, North-Holland.
- WALLIS, K. F., 1972: «Testing for fourth order autocorrelation in quarterly regression equations», *Econometrica*, vol. 40, 617-636.